REPORT DOCUMENTATION PAGE

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE October 15, 2001	3. REPORT TYPE AND DATES COVERED June 24, 1997 - October 18, 2001

4. TITLE AND SUBTITLE
Comparison of Skin Conductance and Skin Resistance Measues for the Detecion of Deception

5. FUNDING NUMBERS
DoDPI97-P-0016

6. AUTHOR(S)
John C. Kircher, Ph.D.; Ted Packard, Ph.D.; Brian G. Bell, Ph.D.; and Paul C. Bernhard, M.S.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Department of Educational Psychology
University of Utah
Salt Lake City, UT 84112

8. PERFORMING ORGANIZATION REPORT NUMBER
DoDPIO2-P-0001

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
DoD Polygraph Institute
7540 Pickens Avenue
Fort Jackson, South Carolina 29207

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
Public Release, distribution unlimited

12a. DISTRIBUTION / AVAILABILITY STATEMENT

13. ABSTRACT *(Maximum 200 words)*
Traditional analog polygraph instruments typically record skin resistance (SR), whereas academic psychophysiologists typically record skin conductance (SC) and have argued that SC is superior to SR. The present study tested if SC or SR is more useful for the detection of deception. 336 participants in a previous experiment (DODPI97-P-0016) were tested about their participation in a mock theft. Half of the participants were guilty of committing the theft, and the other half were innocent. Probable-lie polygraph tests (PL) were administered to half of the innocent and half of the guilty participants, and directed-lie tests (DL) were administered to the remaining participants. Participants were paid $30 and were offered an additional $50 to convince the polygraph examiner of their innocence. Digitized recordings of SC subsequently were transformed to SR. A computer measured the amplitudes and other features of SC responses and SR responses to comparison and relevant test questions.
Both SC and SR were highly diagnostic of truthfulness and deception, and no evidence was found to favor either SC or SR for either PL or DL polygraph tests. Univariate analyses revealed no differences between SC and SR in their ability to discriminate between truthful and deceptive individuals, and multivariate analysis indicated that either measure might be used in combination with other physiological measures to detect deception. However, these conclusions apply only to SR recordings obtained when a constant-voltage is applied to wet electrodes. Additional research would be needed to compare such laboratory-grade measurements of SC and SR to electrodermal activity as it is measured by traditional analog and newer computerized field polygraphs.

14. SUBJECT TERMS

15. NUMBER OF PAGES
19

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT None

DoDPI02-R-0001

Comparison of Skin Conductance and Skin Resistance
Measures for the Detection of Deception

John C. Kircher, Ted Packard, Brian G. Bell and Paul C. Bernhardt
University of Utah, Salt Lake City, Utah 84112

October 18, 2001

Final Report

Project Title: *Comparison of Skin Conductance and Skin Resistance Measures for the Detection of Deception*

Date: 16 April 2001

Revised: 18 October 2001

Submitted by: John C. Kircher, Ph.D. (Principal Investigator)
              Department of Educational Psychology
              University of Utah
              Salt Lake City, UT 84112
              voice: (801) 581-7130
              fax: (801) 581-5566
              email: kircher@gse.utah.edu

              Ted Packard, Ph.D. (Co-Principal Investigator)
              Department of Educational Psychology
              University of Utah
              Salt Lake City, UT 84112
              voice: (801) 581-7130
              fax: (801) 581-5566
              email: kircher@gse.utah.edu

              Paul C. Bernhardt, M.S. (Graduate Assistant)
              Department of Educational Psychology
              University of Utah
              Salt Lake City, UT 84112

              Brian G. Bell, Ph.D. (Research Associate)
              Department of Educational Psychology
              University of Utah
              Salt Lake City, UT 84112

Abstract

Traditional analog polygraph instruments typically record skin resistance (SR), whereas academic psychophysiologists typically record skin conductance (SC) and have argued that SC is superior to SR. The present study tested if SC or SR is more useful for the detection of deception. 336 participants in a previous experiment (DODPI97-P-0016) were tested about their participation in a mock theft. Half of the participants were guilty of committing the theft, and the other half were innocent. Probable-lie polygraph tests (PL) were administered to half of the innocent and half of the guilty participants, and directed-lie tests (DL) were administered to the remaining participants. Participants were paid $30 and were offered an additional $50 to convince the polygraph examiner of their innocence. Digitized recordings of SC subsequently were transformed to SR. A computer measured the amplitudes and other features of SC responses and SR responses to comparison and relevant test questions.

Both SC and SR were highly diagnostic of truthfulness and deception, and no evidence was found to favor either SC or SR for either PL or DL polygraph tests. Univariate analyses revealed no differences between SC and SR in their ability to discriminate between truthful and deceptive individuals, and multivariate analysis indicated that either measure might be used in combination with other physiological measures to detect deception. However, these conclusions apply only to SR recordings obtained when a constant-voltage is applied to wet electrodes. Additional research would be needed to compare such laboratory-grade measurements of SC and SR to electrodermal activity as it is measured by traditional analog and newer computerized field polygraphs.

Introduction

Several methods are currently used by field polygraph examiners to measure electrodermal responses during polygraph examinations. Traditional analog polygraphs typically record skin resistance (SR) from large metal plates placed on the fingertips. These plates develop bias potentials, are subject to movement artifacts, and place high power dissipation requirements on individual sweat glands that can affect sweat gland activity. The recordings from analog instruments often show rapid drops in the baselines, which is a nuisance because the polygraph examiner must frequently re-center the recording pen. To avoid this, some polygraph manufacturers include a filter on their analog instruments that stabilizes the baseline, but it also alters the shape and amplitude of the examinee's SR response to test questions.

By the early 1970s, academic psychophysiologists had abandoned the recording of SR in favor of skin conductance (SC). The advantages in recording SC are that large bias potentials can be avoided, a low constant-voltage circuit (0.5V) is used that has little or no effect on the sweat glands, and the baseline is relatively stable, obviating the need for filtering. In addition, research indicates that SC is related linearly to the number of active sweat glands at the recording site, whereas SR is not (Venables & Christie, 1980). A disadvantage in using the low voltage SC circuit is that it requires the use of wet electrodes. The polygraph examiner must apply a small amount of paste to the electrode before placing it on the skin. The dry metal plates used with traditional analog polygraphs are more convenient because they do not require the use of a conductive electrode gel.

Honts and Barger (1990) compared SC and SR recorded from dry metal plate electrodes attached to four fingers of the same hand. SC was recorded on a traditional analog polygraph with a high, 2.2 V constant-voltage circuit manufactured by Lafayette Instruments (Lafayette, IN). They reported no difference between SC and SR in response amplitudes, although they did find that the SR channel required more pen adjustment than the SC channel during the test. Honts and Barger recommended SC because it requires less adjustment than SR, and it is more closely related to eccrine sweat gland activity.

The failure of Honts and Barger to observe a difference between the amplitudes of SC and SR responses was consistent with results from an earlier study by Boucsein and Hoffman (1979). In contrast to Honts and Barger, Boucsein and Hoffman used laboratory equipment that applied a constant-voltage (0.5V) or a constant current (10 μAmp/cm^2) to wet electrodes. Boucsein and Hoffman found only one difference; the recovery times of SC responses were shorter than were those of SR responses.

The present study evaluated one of several possible differences among methods for assessing subjects' electrodermal responses to test questions. Skin conductance was recorded with the constant-voltage circuit from wet electrodes. Continuous absolute measures of SC were digitized by a computer and subsequently were transformed to SR. Since resistance is the reciprocal of conductance, each calibrated value of conductance (in Siemens) was divided into 1.0 to obtain resistance in ohms. The amplitudes of SC and SR responses were then extracted from the respective response waveforms and compared. Consequently, the present study focused on only the effects of a nonlinear, but monotonic, transformation of SC to SR.

Method

Subjects

Four-hundred-and-seventeen adults were recruited from the general community by newspaper advertisements for a study that examined the effects of the demonstration test on the accuracy of probable lie and directed lie polygraph examinations (DoDPI97-P-0016). The advertisements offered $30 for two hours of participation and the opportunity to earn an additional $50 bonus. Of the 417 individuals, 81 were eliminated from the study. Thirty-three individuals assigned to the guilty condition (16%) declined to participate after they received instructions to commit a simulated theft. Eighteen individuals failed to follow instructions. For example, some individuals did not commit the theft, yet reported for their polygraph test. Others arrived too late or brought a child with them to the lab. Thirteen individuals were dismissed due to health problems, including reports of pain, less than 4 hours sleep, and high blood pressure. Nine individuals assigned to the guilty condition (5%) confessed during the pretest interview. Equipment problems and experimenter errors resulted in the loss of 5 other individuals. The remaining 168 innocent and 168 guilty participants were retained to fill the cells of the design matrix (described below).

The mean age of participants was 30.7 years (SD = 11). Years of education ranged from 9 to 25 (M = 14.3, SD = 2.5). Most participants were Caucasian (87.5%) or Hispanic (5.7%) and either single (53.9%) or married (33.9%). A wide range of occupations was represented, the most common being student (17%), professional (11.9%), sales worker (9.2%), office worker (8.3%), service worker (8.3%), unemployed (7.7%), homemaker (7.7%), and laborer (7.5%).

Design

Participants were assigned randomly to one of 16 groups in a 2 X 2 X 4 factorial design, with equal numbers of male and female participants assigned to each cell. All factors except Sex are represented in the design illustrated in Figure 1.

The first factor, Guilt, had two levels; 168 participants were guilty of committing a mock theft, and the remaining 168 were innocent. The second factor, Test Type, also had two levels; half of the participants were given probable-lie comparison question tests (PL), and half were given directed-lie tests (DL).

Figure 1. *Design of Experiment.*
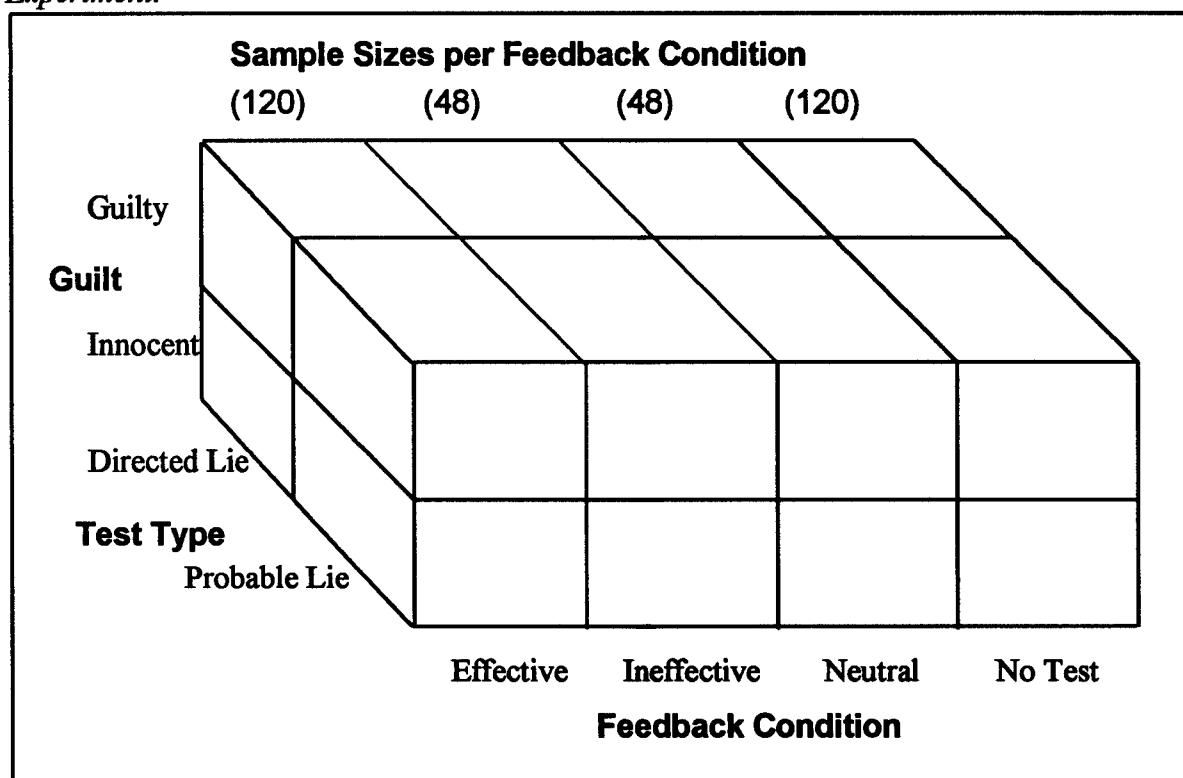
The third factor, Effectiveness Feedback, had four levels. Participants were unevenly distributed over the four levels of the Effectiveness Feedback factor. One group of 120 participants (30 participants in each of the four cells shown on the far left of Figure 1) received the type of feedback commonly provided to subjects in actual field examinations. Prior to their polygraph test, they were asked to select a number and were given a numbers test. They were then told, regardless of the outcome, that they showed their strongest reaction to the number they had chosen. They were also told they would have no problem passing the polygraph test if they were completely truthful to all of the questions (effective-feedback group).

Twelve participants were assigned to each of the four ineffective-feedback cells of the design matrix. Participants who receive ineffective feedback were given a numbers test and told, regardless of the outcome, that they did not react appropriately to the chosen number and that they may not be suitable for examination using a polygraph.

Twelve participants were assigned to each of the four neutral-feedback cells of the matrix. Participants who received neutral feedback were given a numbers test and told that the numbers test would provide an opportunity for the participant to practice answering questions and for the polygraph examiner to adjust the instrument. Participants were given no information about the outcome of the numbers test.

Thirty participants were assigned to each of the four no-test control groups, as illustrated on the far right of Figure 1. Participants in the control groups were not given a numbers test. Otherwise, the pretest procedures were the same as those for all other participants.

To summarize, 120 participants were given the demonstration test and received feedback that the test was effective. Another 48 participants were given a demonstration test and received feedback that the test was ineffective. Another 48 participants were given a demonstration test and received neutral feedback. The remaining 120 participants were not given a demonstration test. Within each level of the Feedback factor, the design was balanced in terms of numbers of guilty and innocent male and female subjects who were given either probable-lie or directed-lie polygraph examinations.

All polygraph tests were administered by two examiners. One examiner was an advanced graduate student in educational psychology. The graduate student (PCB) tested 12 subjects in each of the 16 cells in the design matrix (192 subjects). The remaining 144 subjects were tested by the post-doctoral research associate (BGB).

Procedures

The procedures followed those described elsewhere (Kircher & Raskin, 1988). Prospective participants called a secretary who screened the participants for eligibility and briefly described the experiment and policy for payment. Callers were invited to participate if they met the following criteria: (1) they were between 18 and 65, (2) they were not taking prescription medications, (3) they had never taken a polygraph test, (4) they were fluent in English, and (5) they reported no major medical problems.

Callers who qualified and agreed to participate were given an appointment to report to a room in a building on the campus of the University of Utah. A map and reminder letter were mailed to participants who were scheduled more than a couple of days prior to their appointment. The participant also was called or sent an electronic mail message as a reminder the evening before their appointment.

An envelope addressed to the participant was taped to the door of the room to which the participant reported. Instructions within the envelope directed the participant to enter the room, close the door, read and sign an informed consent form, complete a brief questionnaire, and then play a cassette recorder that presented their instructions over earphones.

Guilty participants heard tape-recorded instructions to commit a mock theft of a $20 bill from a wallet in a purse in a secretary's desk. Participants went to a secretary's office where they asked the secretary where Dr. Mitchell's office was located. The secretary was actually a confederate in the experiment. The secretary responded that there was no Dr. Mitchell in the department, and the participant left the secretary's office. The participant then watched for the secretary to leave the office unattended (1-3 minutes), entered the office, searched the desk for the purse, took the wallet from the purse, took the $20 bill from the wallet, and concealed the $20 on their person. Participants then reported to a waiting room where they waited for the polygraph examiner. Guilty participants were also instructed to prepare an alibi in case they were caught in the office. Innocent participants listened to a general description of the crime, left the area for 15 minutes, and then reported to the waiting room.

All participants were told that they would be given a polygraph test by a polygraph expert who would not know if they had committed the theft. They were told that the examiner would use a computer to assist in the analysis of their polygraph charts. They were told that if they convinced

the polygraph examiner of their innocence, they would receive a total of $80. They also were told that if they failed to convince the examiner of their innocence, they would receive only $30.

When the polygraph examiner went to the waiting room, he asked participants to use the restroom and wash their hands. When they returned from the restroom, they were escorted to the laboratory and asked to sit in the examinee's chair. The session was video and audio taped using a camera mounted high on the wall in front of the participant.

Standard field polygraph procedures were used. The polygraph examiner asked participants' about their prior experiences with the polygraph. The examiner then asked participants to sign a Polygraph Informed Consent form. The examiner then obtained biographical information and asked questions about their health. Participants who had less than 4 hours of sleep, were experiencing pain, or indicated that they had recently taken stimulant or depressant drugs (prescription or otherwise) were not tested, were paid for their partial participation, and released. The sensors were attached and adjusted to ensure adequate recordings. The examiner then described the role of the autonomic nervous system in the detection of deception. The demonstration test was then conducted if the participant was in a neutral, effective, or ineffective feedback condition.

After the demonstration test, the examiner reviewed the appropriate set of questions with the participant. The test questions for participants assigned to the probable-lie condition were as follows:

(Outside Issue)	1.	Do you understand that I will ask only the questions we have discussed?
(Sacrifice Relevant)	2.	Do you intend to answer truthfully all of the questions about the theft of the $20?
(Neutral)	3.	Do you live in the United States?
(Probable-lie)	4.	Before the age of __, did you ever take something that didn't belong to you?
(Relevant)	5.	Did you take that $20 from the purse?
(Neutral)	6.	Is today ___?
(Probable-lie)	7.	During the first __ years of your life, did you ever do anything that was dishonest or illegal?
(Relevant)	8.	Did you take that $20?
(Neutral)	9.	Is your first name ___?
(Probable-lie)	10.	Between the ages of __ and __, did you ever lie to get out of trouble?
(Relevant)	11.	Do you have that $20 with you now?

Relevant questions that pertained to the theft and the sacrifice relevant question were reviewed first, probable-lie or directed-lie comparison questions were reviewed next, and the neutral and outside issue questions were reviewed last. When the examiner introduced the probable-lie questions, the examiner indicated that the questions were about the examinee's basic honesty, and their purpose was to determine if they were the type of person who would take something then lie about it. If the participant answered "Yes" to a probable-lie question, the question was reworded slightly to elicit a "No" response from the participant; e.g., "Other than what you told me, before the age of __, did you ever take something that didn't belong to you?

The test questions for participants assigned to the directed-lie condition were the same as those presented to participants in the probable-lie condition, except that the probable-lie questions in positions 4, 7, and 10 were replaced with the following directed-lie questions.

(Directed-lie) 4. In your entire life, did you ever tell even one lie?
(Directed-lie) 7. Have you ever broken a rule or regulation?
(Directed-lie) 10. Did you ever make a mistake?

Participants were told that it was very important that they appear to be lying to the directed-lie questions. The examiner told the participant that he would not want to make a mistake and conclude that they had lied if they were in fact telling the truth, simply because they did not appear to be lying on these questions.

The probable-lie or directed-lie test was then administered. The interval between question onsets was a minimum of 25 s, and the interval between charts was between one and three minutes. After the first chart, probable-lie participants were asked if there were any problems with any of the questions. After the second chart, they were asked if they felt anything unusual when they were asked one of the probable-lie questions. Conversely, directed-lie participants were asked after each chart if they were lying to the directed-lie questions and if they felt any differently when they lied. These procedures were designed to draw the participant's attention to the comparison questions and reduce the risk of false positive errors.

The question sequence was presented five times, which provided five charts of data. Neutral and comparison questions were rotated over repeated presentations of the question sequence such that each relevant question was preceded by each neutral and each comparison question at least once. The order of presentation of the questions was not reviewed with the participant in advance of collecting the physiological data.

At the conclusion of the test, the sensors were removed, and the subject was asked to complete post-test questionnaires. After minor editing of obvious movement artifacts from the physiological data, the probability that the participant was truthful was then determined using computer algorithms described elsewhere (Kircher & Raskin, 1988). If the probability of truthfulness exceeded 0.70, the participant was diagnosed as truthful and paid $80, $30 for their time and a $50 bonus. If the probability of truthfulness was less than 0.30, the participant was diagnosed as deceptive. If the probability of truthfulness was between 0.30 and 0.70, the test was considered inconclusive. If the outcome was deceptive or inconclusive, the participant was paid only $30. The participant was then debriefed and released.

Apparatus

The CPS-LAB system (Scientific Assessment Technologies, Salt Lake City, UT) was used to configure the data collection hardware, specify storage rates for the physiological signals, and build automated data collection protocols. CPS-LAB also was used to collect, edit, and score the physiological data.

The physiological data acquisition subsystem (PDAS) of CPS-LAB generated analog signals for skin conductance, thoracic and abdominal respiration, cardiograph, finger pulse amplitude, skin

potential, and cardiotachometer. In addition, calibrated analog output from a Ohmeda 2300 Blood Pressure Monitor was routed to a general-purpose coupler on the PDAS. Each of the eight analog signals was digitized at 1000 Hz with a Metrabyte DAS 16F analog-to-digital converter installed in a 50 MHz IBM-PC compatible 486 computer.

Skin conductance was obtained by applying a constant voltage of 0.5V to two Beckman 10mm Ag-AgCl electrodes filled with .05 M NaCl in a Unibase medium. The electrodes were attached with double-sided-adhesive collars to the middle phalanx of the ring and smallest fingers of the left hand. The signal was recorded DC-coupled with a 2-pole, low-pass filter, fc = 6 Hz.

Respiration was recorded from two Hg strain gauges secured with Velcro straps around the upper chest and abdomen, just below the ribcage. The strain gauge changed in resistance as the subject breathed. Resistance changes were recorded DC-coupled with a 2-pole, low-pass filter, f_c = 8.8 Hz.

The cardiograph was recorded from a blood pressure cuff wrapped around the right upper arm and inflated to 55-60 mm Hg at the beginning of each chart. The cuff was connected by rubber tubing to a pressure transducer in the PDAS. The output from the pressure transducer was amplified and recorded DC-coupled with a 2-pole, low-pass filter, fc = 8.8 Hz.

The procedures for measuring finger pulse, electrocardiogram, skin potential, and blood pressure are described in detail in another report (Kircher, Packard, Bell, & Bernhardt, 2001).

The 1000 Hz samples for each channel were reduced prior to storing them on the hard disk by averaging successive sample points. Electrodermal and respiration channels were stored at 10 Hz. Cardiograph signals were stored at 100 Hz.

Calibration Procedures

To test for differences between skin conductance and skin resistance measures, it was necessary to convert the raw data provided by the analog-to-digital converter to absolute units of skin conductance and, subsequently, resistance. To measure skin conductance, a separate multiple regression equation was developed for each of six possible gain settings on the PDAS. Each equation predicted known conductances from the offset on the front panel, internal PDAS digital-to-analog (DAC) offset settings, and observed analog-to-digital converter values. The conductance values used to calibrate the instrument ranged from 1 µSiemen to 50 µSiemens. External (front panel) and internal (DAC) offsets were also systematically varied to ensure that the resulting equation would work for any combination of gain and offset settings. Each equation accounted for over 99.8% of the variance in known inputs.

Since resistance (R) is the reciprocal of conductance (G), skin resistance was obtained by inverting the scaled skin conductance signal prior to extracting measurements of response amplitude, i.e., $R = 1 / G$.

Feature Extraction

Feature extraction was accomplished with the CPS-LAB computer program. The features

extracted from each physiological channel were those that have been found in previous investigations to be optimal for the prediction of deception in a laboratory mock crime study.

Skin Conductance and Skin Resistance. For each comparison and relevant question, the peak amplitude of SC and SR was extracted from a 20-second interval that began at question onset. Although the dependent variables of greatest interest were the amplitudes of SC and SR responses, additional response features were extracted to test if there was any systematic difference between SC and SR responses in terms of their overall diagnosticity. Figure 2 illustrates the measurement of response amplitude and selected additional features. Full recovery time, duration to full recovery, full recovery rate, and area are not illustrated in Figure 2, but they were also measured. The point of full recovery occurred when the recovery limb of the response reached the level at response onset. The response shown in Figure 2 did not recover to the level at response onset before the end of the 20-second scoring window. When the response did not recover completely, the point of full recovery was taken as the end of the 20-second scoring window.

Figure 2. Feature Extraction
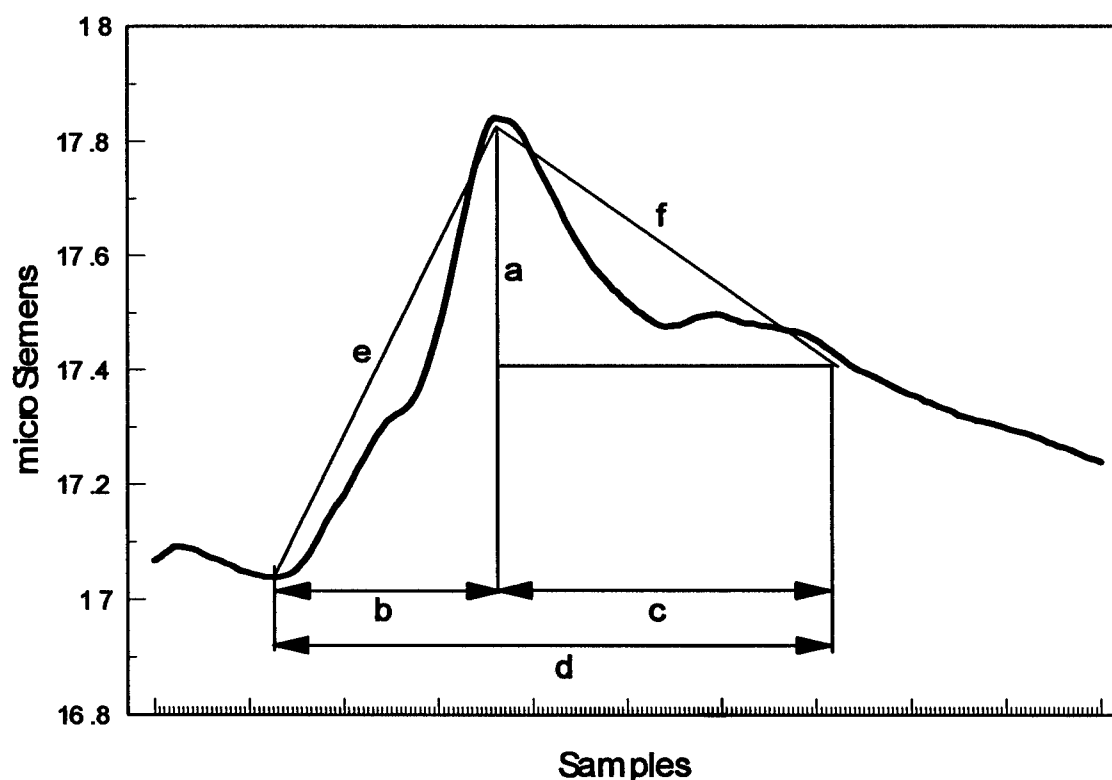
a. Amplitude
a. Rise time
b. Half-recovery time
c. Duration to half recovery (b+c)
d. Rise rate (a / b)
e. Half-recovery rate (a/2 / c)

Cardiograph. The times and levels of systolic and diastolic points were identified in the cardiograph signal and used to create second-by-second systolic and diastolic response curves

(Kircher & Raskin, 1988). A trend line was computed by calculating the mean of the systolic and diastolic points for each second. The maximum increase in the trend line (response amplitude) was extracted from the 20-second interval that followed question onset.

Thoracic and Abdominal Respiration. Thoracic and abdominal respiration excursion was measured for a period of 10 seconds following question onset. Since each respiration channel was stored at 10 Hz, 100 samples showed the respiration activity for the 10-second interval. Respiration excursion was the sum of 99 absolute differences between adjacent samples.

Feature Standardization

The present analyses were limited to the first three charts of physiological data. Each chart contained three comparison questions and three relevant questions. The six questions on each of the three charts provided 18 repeated measures of each physiological feature. The set of 18 measurements of a given feature were transformed to standard scores (z-scores) within the subject. A mean z-score was calculated for the comparison questions and another mean z-score was calculated for the relevant questions. The difference between the means for comparison and relevant questions provided an overall index of differential reactivity to comparison and relevant questions.

For SC amplitude, SR amplitude, and cardiograph amplitude, a large measured value was interpreted as a strong physiological reaction to the question. For respiration excursion, a relatively small measured response was indicative of strong respiration suppression. To establish a common direction for predicted effects, the sign of the difference between comparison and relevant questions was reversed for respiration excursion. Thus, for all measures, stronger reactions to comparison questions resulted in positive difference scores, and stronger reactions to relevant questions resulted in negative difference scores. A single composite measure of differential respiration suppression was then obtained by computing the mean of the difference scores for thoracic and abdominal respiration excursion.

Results

Preliminary Analyses

Treatment-Related Attrition. Thirty-three individuals assigned to the guilty condition (15%) refused to participate after they had received their tape-recorded instructions, whereas none of the innocent subjects declined to participate. Consequently, individuals who agreed to commit the mock crime may have been sampled from a population that differed in certain respects from the more general population from which innocent participants were drawn. For example, participants in the guilty condition on average may have been older or less anxious than were participants in the innocent condition. Preliminary tests were conducted to explore the possibility that guilty and innocent participants differed on measures of marital status, ethnicity, occupation, age, education, hours of sleep, the Marlowe-Crowne scale (Crown & Marlowe, 1964), Rotter Trust Scale (Rotter, 1967), and two anxiety scales (Spielberger, Gorsuch, Lushene, Vagg, & Jacobs, 1983). The guilty and innocent participants who completed the experiment did not differ significantly on any of the demographic or personality measures.

Effects of Guilt, Feedback, and Test Type. A split-plot ANOVA was performed to test if differences between SC amplitude and SR amplitude varied across the Guilt, Feedback, and Test Type facets of the design. Method of measurement was the within-subjects factor with two levels (SC amplitude versus SR amplitude). Guilt, Feedback (effective, neutral, ineffective, and no test), and Test Type (PL versus DL) were between-group factors. If differences between SC amplitude and SR amplitude varied over levels of Feedback or Test Type, it might not be possible to conclude that one electrodermal measure was generally superior to the other.

The results of the split-plot ANOVA revealed no main effect of Method nor was there any significant 2-, 3-, or 4-way interaction of Method with Guilt, Feedback, or Test Type. The lack of a significant Guilt X Method interaction suggests that SC amplitude and SR amplitude were similarly useful for discriminating between guilty and innocent participants when the data were pooled across Test Types and Feedback conditions. The lack of any higher-order interactions that included Guilt and Method as factors suggests that the *difference* between SC amplitude and SR amplitude in their ability to discriminate between guilty and innocent subjects (Guilt X Method interaction) did not depend on the Test Type or Feedback.

Univariate Analyses

The objective of the present study was to determine if there is an advantage in using SC or SR for the detection of deception. To that end, the preliminary analyses indicated that it is reasonable to pool the data across PL and DL tests and across Feedback conditions and to capitalize on the large sample size. Therefore, comparisons based on all 336 participants are presented below. However, the preliminary ANOVA also revealed small but significant Guilt X Test Type, $F(1, 320) = 5.49$, $\eta^2 = .02$, and Guilt X Feedback X Test Type interactions effects, $F(3, 320) = 2.73$, $\eta^2 = .03$, on both electrodermal measures. Since only the effective feedback condition was representative of current field practice, results also are reported separately for the PL and DL effective feedback conditions.

Table 1 presents the means of the differences between standardized responses to comparison and relevant questions for the SC amplitude and SR amplitude measures. The means are presented separately for the guilty and innocent participants in the entire sample and for the PL and DL effective feedback subsamples. Table 1 also shows the point-biserial correlation between the index of differential reactivity and a dichotomous criterion variable that distinguished between the guilty (coded 0) and innocent participants (coded 1) in a sample. The point-biserial correlation is a measure of diagnostic validity; it indicates the extent to which the physiological measure discriminated between the innocent and guilty groups.

Overall, the validity coefficients for SC and SR were lower in the entire sample ($M_r = .61$) than in either of the effective feedback subsamples (Probable-Lie $M_r = .72$; Directed Lie $M_r = .64$). ANOVAs revealed significant Guilt X Feedback interactions for SC ($F(1, 328) = 4.59$, $p < .05$) and SR ($F(1, 328) = 4.04$, $p < .05$) when the effective feedback conditions jointly were compared to the groups that did not receive effective feedback. These findings suggest that a preliminary demonstration test and effective feedback enhance the accuracy of polygraph tests. These findings are discussed in detail in another report (Kircher, Packard, Bell, & Bernhardt, 2001.

Table 1. Means (Standard Deviations) and Point-Biserial Correlations (r_{pb}) for SC Amplitude and SR Amplitude

	Innocent		Guilty		r_{pb}
Entire Sample (N=336)					
SC Amplitude	.46	(.62)	-.50	(.64)	.61**
SR Amplitude	.44	(.62)	-.52	(.64)	.61**
Probable-Lie Effective Feedback (n=60)					
SC Amplitude	.42	(.62)	-.84	(.60)	.72**
SR Amplitude	.40	(.62)	-.74	(.62)	.72**
Directed-Lie Effective Feedback (n=60)					
SC Amplitude	.42	(.70)	-.62	(.58)	.63**
SR Amplitude	.40	(.66)	-.64	(.62)	.64**

** $p < .01$

There was a tendency for the electrodermal measures to be more diagnostic for effective feedback participants who received PL tests ($M_r = .72$) than DL tests ($M_r = .64$), but the difference was not significant for SC amplitude or for SR amplitude. Of greater interest in the present study were comparisons of SC and SR measures within treatment conditions. The validity coefficients for SC amplitude and SR amplitude were virtually identical for the entire sample and within the PL and DL effective feedback conditions. A separate z-test for correlated correlations (McNemar, 1969) was performed for each sample. The z-tests confirmed that there were no significant differences between the validity coefficients for SC and SR measures for the entire sample (.61 versus .61), the PL effective feedback group (.72 versus .72), or the DL effective feedback group (.63 versus .64).

Numerical evaluations of electrodermal responses depend primarily on measures of response amplitude (Bell et al., 1999; Swinford, 1999), and the only feature extracted from the electrodermal response by our CPS computer program is peak amplitude. However, other aspects of the electrodermal response, such as duration and number of responses, are used by numerical evaluators and might be used by computer programs in the present or future. Table 2 presents the validity coefficients for selected features of the SC and SR waveforms, including the amplitude measures presented previously. Since the SC and SR waveforms are monotonically related, there was always an equal number of SC and SR responses. Although there was no possibility that the number of SC and SR responses would differ, the point-biserial correlations for number of responses are reported for completeness.

Table 2. Point-biserial Correlations with the Guilt/Innocence Criterion for Features Extracted from SC and SR Waveforms

Feature	Entire Sample (N=336)		Probable-lie Effective Feedback (n=60)		Directed Lie Effective Feedback (n=60)	
	SC	SR	SC	SR	SC	SR
Amplitude	.61	.61	.72	.72	.63	.64
Rise time	.23	.23	.25	.25	.37	.37
Half-recovery time	.48	.48	.55	.56	.43	.44
Full-recovery time	.57	.57	.64	.64	.61	.61
Duration to half-recovery	.37	.38	.40	.41	.45	.45
Duration to full-recovery	.57	.57	.65	.65	.62	.61
Number of responses	.25	.25	.39	.39	.30	.30
Area to half-recovery	.55	.55	.68	.66	.54	.53
Area to full-recovery	.58	.58	.71	.69	.56	.56
Rise rate	.57	.58	.67	.67	.63	.63
Recovery rate to half-rec	.43	.43	.57	.58	.52	.52
Recovery rate to full-rec	.36	.36	.55	.55	.35	.35

Note: All of the correlation coefficients in Table 2 were statistically greater than zero, $p < .05$.

The results in Table 2 are consistent with those obtained for SC amplitude and SR amplitude. There was virtually no difference in diagnostic validity of various aspects of the SC and SR response waveforms.

Boucsein and Hoffman (1979) found no difference between the amplitudes of SC and SR responses, but they did report that the recovery times of SC responses were shorter than were those of SR responses. To assess the reliability of that finding, the mean of raw measurements of SC half recovery time across all comparison and relevant questions was obtained for each participant. Another mean was obtained for SR half recovery times. Method of measurement (SC versus SR) was treated as a repeated measure in an ANOVA with Guilt, Test Type, and Feedback as between-group factors. Consistent with the results reported by Boucsein and Hoffman, SC half recovery times (M = 3.77 sec) were significantly shorter than the half recovery times of SR responses (M = 4.08 sec), $F (1, 320) = 373.7$, $p < .01$.

Multivariate Analyses

It is conceivable that one variable may be correlated with the criterion as highly as or even less highly than another, and yet it produces higher hit rates when combined with other physiological measures. To evaluate this possibility, a discriminant function was created to classify cases that included SC amplitude along with respiration excursion and cardiograph baseline increases. Another discriminant function was created that included SR amplitude as well as the respiration and cardiograph measures. Cases were classified as truthful if the probability of truthfulness exceeded 0.70, deceptive if the probability of truthfulness was less than 0.30, and inconclusive if the probability was between 0.70 and 0.30. Percent correct, wrong, and inconclusive (?) that resulted from this decision rule are presented in Table 3.

Table 3. Percent Outcomes for Discriminant Functions that Combine SC Amplitude or SR Amplitude with Cardiograph and Respiration Measures

	Innocent				Guilty			
	Correct	Wrong	?	Correct Decisions	Correct	Wrong	?	Correct Decisions
Entire Sample (N=336)								
SC Amplitude	63.7	9.5	26.8	87.0	63.1	10.1	26.8	86.2
SR Amplitude	62.5	9.5	28.0	86.8	64.9	10.7	24.4	85.8
Probable-Lie Effective Feedback (n=60)								
SC Amplitude	90.0	6.7	3.3	93.1	86.7	6.7	6.7	92.9
SR Amplitude	90.0	3.3	6.7	96.4	83.3	6.7	10.0	92.6
Directed-Lie Effective Feedback (n=60)								
SC Amplitude	70.0	16.7	13.3	80.8	73.3	10.0	16.7	88.0
SR Amplitude	76.7	16.7	6.7	82.1	70.0	10.0	20.0	87.6

In the entire sample, and within each effective feedback condition, there was little difference between outcomes of discriminant functions that included SC or SR responses. A post hoc examination of Table 3 suggested that the inconclusive rate appeared higher in the entire sample than in either of the effective feedback subsamples. To test this hypothesis, subjects who received effective feedback were dropped from the entire sample, and frequencies of correct decision, wrong decision, and inconclusive were calculated for the remaining three groups. Another group was formed by pooling the results from the PL and DL subjects who received effective feedback. A chi-square test was then conducted to test if the distribution of outcomes (correct, wrong, and inconclusive) depended on whether or not the subject had received effective feedback. The test confirmed that the presentation of effective feedback affected outcomes, $X^2(2) = 27.8$, $p < .01$. The effect was due primarily to the relatively low number of inconclusive cases in the group that received effective feedback. Another chi-square test compared outcomes from the PL and DL effective feedback groups. Although there appeared to be some advantage in using the PL test, the difference was not statistically significant, $X^2(2) = 5.4$, $p < .07$.

Discussion

The results of our comparisons of SC and SR are consistent with those reported by Boucsein and Hoffman (1979) and by Honts and Barger (1990). We observed no differences between SC amplitude and SR amplitude measurements across a wide range of treatment conditions.

Boucsein and Hoffman also reported that the recovery time of SC responses was shorter than the recovery time of SR responses. We replicated that finding as well. However, within-subject comparisons of SC and SR recovery times revealed no differences in their ability to distinguish between truthful and deceptive individuals.

In the present study, SC recordings were transformed mathematically to SR prior to extracting features from them. Although this transformation was nonlinear, it had no discernable effect on any of the 12 response characteristics measured in the present study. Visual comparisons of individual SC and SR responses by several participants suggested that within the range of measurements for an individual, the inverse transformation from SC to SR was essentially linear. The greatest observed difference in the shape of the SC and SR waveforms occurred when the basal level of SC was low (e.g., 1 µS) and the reactions were large relative to the basal level (e.g., 0.5 µS). Even then, the transformation produced an SR waveform that appeared very similar to the SC waveform on most dimensions. The results of visual inspection and computer analysis were consistent; SC and SR response waveforms were virtually identical.

The consequences of failing a polygraph test administered during an actual criminal investigation are usually much greater than those associated with failing a test in a laboratory experiment. On average, tonic levels of arousal may be greater in the field than in the laboratory. Given that SC and SR responses appear more similar at elevated levels of tonic activity, it seems unlikely that significant differences between SC and SR measures would emerge in the field. Nevertheless, it should be noted that this was a mock crime experiment and the present findings may not generalize to the field.

The present study evaluated only one of several possible differences in methods for measuring participants' electrodermal responses to test questions. SR was based on measurements obtained with a constant low-voltage source, but it is usually obtained from a constant current source. The design of the present study did not permit a test for possible differences in recording techniques. In terms of decision accuracy, prior research suggests there would be no particular advantage in choosing a constant-voltage over a constant-current circuit for polygraph testing (Boucsein & Hoffman, 1979; Honts & Barger, 1990). However, we agree with Honts and Barger (1990) that the constant-voltage SC circuit is preferred. We agree because it produces measures that are related linearly to the number of active eccrine sweat glands, it produces a stable baseline that does not require a high pass filter, and it is consistent with accepted scientific practice. Boucsein and Hoffman (1979) used laboratory equipment with wet nonpolarizing electrodes, whereas Honts and Barger (1990) used an analog field polygraph with dry metal plate electrodes. Since neither study compared laboratory instrumentation to field polygraphs, we do not yet know if laboratory equipment and techniques yield electrodermal measures that are more diagnostic of deception than those from traditional analog polygraph instruments.

Another unanswered question concerns the differences among computerized polygraph systems in their methods for recording electrodermal activity. The Computerized Polygraph System (CPS; Stoelting Company, Wood Dale, IL), the Axciton system (Axciton Systems, Houston, TX), and the Lafayette system (Lafayette Instruments, Lafayette, IN) are currently used by field polygraph examiners. Only the CPS system records SC from wet electrodes with a constant-voltage circuit and meets guidelines for recording electrodermal activity established by the scientific community (Fowles, Christie, Edelberg, Grings, Lykken, & Venables, 1981). The other systems use traditional dry metal plates as electrodes. However, in contrast to traditional analog

instruments, the Axciton and Lafayette computerized polygraphs do not measure SR. In his tests of the three computerized polygraph systems, Cestaro (1998) found that the signals generated by Axciton and Lafayette computer systems did not accurately reproduce known changes in resistance or conductance. Although there are limitations to the traditional methods for recording SR, at least there is a more or less direct (monotonic) relationship between SR and the activity of the eccrine sweat glands (Venables & Christie, 1980). In light of Cestaro's findings, the same cannot be said of the electrodermal signals generated by the Axciton and Lafayette computerized polygraphs. Therefore, it is also important to determine if the electrodermal measures from laboratory-grade polygraph instruments, such as CPS, are more useful for detecting deception than those provided by other computerized polygraph instruments.

References

Boucsein, W. & Hoffman, G. (1979). A direct comparison of skin conductance and skin resistance methods. *Psychophysiology, 16,* 66-70.

Cestaro, V. (1998). Memorandum for Record: Laboratory tests performed on the electrodermal activity (EDA) channels of various polygraph instruments. Report on project DoDPI98-P-0003 to the U. S. Department of Defense Polygraph Institute, Ft. McCellan, AL.

Crowne, D.P., & Marlowe, D. (1964). *The approval motive: Studies in evaluative dependence.* New York: Wiley.

Honts, C. R. & Barger, S. D. (1990). A comparison of the relative utility of skin conductance and skin resistance couplers for the measurement of electrodermal activity. *Polygraph, 19,* 199-207.

Fowles, D. C., Christie, M. J., Edelberg, R., Grings,W. W., Lykken, D. T., & Venables, P. H. (1981). Publication recommendations for electrodermal measurements. *Psychophysiology, 16,* 66-70.

Kircher, J. C., Packard, T., Bell, B. G. & Bernhardt, P. C., (2001). Effects of Prior Demonstrations of Polygraph Accuracy on Outcomes of Probable-Lie and Directed-lie Polygraph Tests. Final report to the U. S. Department of Defense Polygraph Institute, Ft. Jackson, SC. Salt Lake City: University of Utah, Department of Educational Psychology.

Kircher, J.C. & Raskin, D.C. (1988). Human versus computerized evaluations of polygraph data in a laboratory setting. *Journal of Applied Psychology, 73,* 291-302.

Kircher, J. C., Woltz, D. J., Bell, B. G., & Bernhardt, P. C. (1998). Effects of audiovisual presentations of test questions during relevant-irrelevant polygraph examinations and new measures. Final Report to the U. S. Central Intelligence Agency. Salt Lake City: University of Utah, Department of Educational Psychology.

Rotter, J.B. (1967). A new scale for the measurement of interpersonal trust. *Journal of Personality, 35,* 651-665.

Spielberger, C.D., Gorsuch, R.L., Lushene, R., Vagg, P.R., & Jacobs, G.A. (1983). *Manual for the State-Trait Anxiety Inventory.* Palo Alto, CA.: Consulting Psycholgists Press.

Venables, P.H. & Christie, M.J. (1980). Electrodermal activity. In I. Martin & P.H. Venables (Eds.) *Techniques in Psychophysiology.* Chichester: Wiley.